

# Shashi Budati

 [shashibudati@gmail.com](mailto:shashibudati@gmail.com) | [LinkedIn](#) |  +91-9493411471 |  Hyderabad

I am a passionate and detail-oriented **Full Stack Software Engineer**, excited about learning and incubating new technologies. I am looking to obtain a position where I can work on challenging and impactful projects.

## Work Experience

---

- Experienced with owning end-to-end software development processes including– gathering requirements from product meets, solutioning, building services from scratch, reviewing code, documenting religiously, setting up pipelines to meet quality-gates, handling release cadence, managing deployments and configuration, handling on-call, monitoring analytical tools, providing technical support, etc.
- Have experience working effectively on multiple scopes between teams, by providing feedback on ADRs, processes.
- Have taken recruitment calls actively, maintained team-specific handbooks for hassle-free onboarding of new-hires, and mentored associate developers across the org.

**SDE III, Engineering @ Setu (PineLabs)**, Bangalore

**May 2021 – Current**

### Data Insights, Bank Statement Analysis

**Oct 2022 – Current**

Built a data-analytics project from scratch that processes bank statements and Account Aggregator FIU data to provide accurate insights for underwriting, income verification, and expense categorizations.

Achieved ~86% precision and recall for classification tasks.

- Wore multiple hats as a founding engineer, handling data annotation, data engineering, data science, ML-Ops.
- Developed backend pipelines, processes, internal platforms and solutions to scale team operations effectively.
- Developed a rule-based tokenizer/parser using **Antlr**. Have set up an efficient process to find distinct lexical patterns from large datasets, annotate recurring patterns for NER and generate domain-specific tokenizers.
- Developed and deployed classification models based on keyword-matching and embedding-similarity. Used **pyahocorasick**, **Word2Vec**, **spaCy**.
- Post-processed (finetuning) and deployed foundational **LLMs** for running generative tasks on financial data. Used **axolotl** fine-tuning, **vLLM** serving, **Slurm** Workload Manager.
- Developed a configurable pipeline to run inference of combinations of models, optimizing for precision / latency / cost as required for different use-cases or tasks.

## Lending

**May 2021 – Oct 2022**

- Enabled NTB & ETB lending on WhatsApp, by building a couple of bots integrated with 2 banks. Facilitated quick & complete loan application process including integrations for lead generation, underwriting with financial statement analysis, KYC, e-Sign workflows, e-NACHs registrations, etc to loan disbursals.
- Developed a common bot-framework to extend the above use-case, to integrate with different messaging channels and different banks.
- Built a lending dashboard as part of our common console (The Bridge), for bank employees and partner organizations to view and partially-assist leads generated via the bots. Moved design, modeling and authn/z discussions for revamping the console to fit broader use-cases.
- Worked on POC for a lending platform- a marketplace for lenders to get integrated and provide embedded lending, BNPL services across different channels.

## Platform

- Developed and presented an open Keycloak plug-in that supports OTP based login-flows for different protocols including Authorization Code Grant (OAuth2, browser-based) and Direct Grant (OpenIDConnect, API-based).
- Have setup internal platform like Argilla, JupyterLab, DVC repositories, MLFlow tracking server etc to facilitate development of ML / AI models

## Member Technical Staff @ Zoho Corporation (ManageEngine), Chennai May 2017 – May 2021

Was Directly-Responsible-Individual for integrating Threat-Intelligence into existing log product features such as reports, alerts, and compliance.

### Log360Cloud Feb 2019 – May 2021

- Threat Analytics – Built a system comprising a Cassandra cluster storing IP/URL reputation data, and a worker service to analyze uploaded logs using that data.

### EventLog Analyzer, Log360 (On-premises) Jan 2018 - Feb 2019

- Benchmarked different data-stores performances for inserting and searching millions of IP ranges. Enhanced product efficiency by using MapDB's off-heap navigable data-structures, supporting real-time event-log processing.
- Added integration support for STIX/TAXII servers, and used the polled feeds for generating threat intelligence reports and alerts. Adding offline agent-based distribution support to comply with network firewall requirements of our customers.
- Built a Hadoop DFS-based service to integrate and store third-party feeds, and re-distribute them to authenticated on-premises clients and internal micro-services.

### SharePoint Manager Plus Jun – Dec 2017

- Archive File Search – Archived older data into scalable full-text search engine (Lucene) to provide Business intelligence reports without the performance issues of a single-shard embedded database. Handled sharding, concurrency, etc to support Solr / Elasticsearch features in the on-premises setup.

## Personal Projects

---

### Navigable Stick for the blind Dec 2015 - May 2016

- Designed a walking stick giving directional haptic feedback for guiding navigation. Developed an Android application integrated with Google Maps API, which can take voice commands to start navigation. The stick's MicroProcessor transmits gyroscopic sensor information to the application to coordinate directions.
- Presented the application at JHub Excite Workshop 2016

## Education

---

### Computer Science and Engineering @ JNTUH-CES 2013-2017

#### 3D Modeling using Android July - Dec 2016

- Developed an Android application providing small-range (non-aerial) 3D mesh model in OpenGL, from camera feeds of mobiles in WiFi-P2P network, using Computer Vision and PhotoGrammetry techniques.

#### Quality-of-Experience Attribute Mining, using Natural Language Processing techniques Jan - May 2017

- Developed a Python application providing a quantified score for online reviews, by crawling and processing them using Natural Language Toolkit (NLTK) for Python. The quantified scores could then be used for automation and better user experience.

## Knowledge Areas

---

- Languages – **Python, JavaScript, Java**. Data structures.
- ZOHO Cloud Framework, **AWS** platform – Worked with EC2, ECS (Docker), RDS, S3, Glue, Athena, etc
  - Worked with terraform and custom tooling to manage infrastructure
- RDBMS with **PostgreSQL**, MySQL, and JPA/ORM using Hibernate, SQLAlchemy, Alembic. NoSQL using **Cassandra**. Full-text using Lucene, **ElasticSearch**.
  - Navigable Maps using [MapDB](#). Serialization using [Kryo](#).
- Data engineering and analytics - Worked on NLP, ML requirements
- Compiler theory - Antlr4 language parser generator
- Cyber Threat Intelligence – Mitre [Att&Ck](#), [STIX](#) / [TAXII](#)
- Caching Layer using **Redis**. Queuing Layer using [Celery](#) + **SQS**
- Migrated [Redash](#) to [Metabase](#). Worked with [Grafana](#)
- Migrated [Flask](#)-based projects to [FastAPI](#), [Poetry](#). Worked with [Jupyter](#).
- Single Page Web Applications, Front-end MVVM frameworks – [Ember.js](#), [Vue.js](#), [React.js](#)